Recognition of fricatives in normal hearing and simulated hearing-impaired conditions

Examination Number: B001855

Master of Science
in Developmental Linguistics
Linguistics and English Language
School of Philosophy, Psychology and Language Sciences
University of Edinburgh
2011

Acknowledgment

I would like to thank Dr Mits Ota and Dr Maria Wolters for supervising my work.

Special thanks go to my friends and family for encouraging me to go ahead with my project, and to Andrew Buskell whose stimulating suggestions and encouragement helped me structure and organise my paper. I am also deeply indebted to Libby Cunningham and Siôn Dominic Philpott-Morgan for correcting style and grammar of my paper and offering suggestions for improvement.

Finally I would like to express my sincerest gratitude to all those who gave me the possibility to complete the present study.

Abstract

Previous studies on consonant recognition and simulated hearing impairment have shown that consonants most often get confused in noise conditions (Miller and Nicely, 1955, Maniwa et al., 2008). This study follows these previous studies by investigating how our linguistic perception of fricatives can be affected by various auditory conditions. In particular, the four voiceless fricatives /f/, /θ/, /s/ and /ʃ/ have been embedded in the words *fin*, *thin*, *sin* and *shin* and recorded in the carrier sentence "I say ______ now" by a female and a male speaker of Scottish English. The 32 sentences were presented to students at the University of Edinburgh twice in random order in a normal hearing and a simulated hearing-impaired condition with and without noise. The sentences were played over a headset with the 4 target words simultaneously presented on the computer screen. The listeners' task was to select the word they heard. Analyses were conducted on talker's voice, normal and impairment conditions regarding the listeners' results. Findings showed that the talker's voice, noise and reaction time all had an effect on the listener's recognition of fricatives. As expected the noise condition produced higher incorrect results. By trying to replicate earlier studies in a different language environment, some of the results confirmed the findings by Miller and Nicely (1955) and Maniwa et al. (2008) with regard to consonant recognition and confusion. The study also revealed that noise affected the recognition more than the impairment condition and that fricatives in the male voice condition were overall recognised better.

Table of Content

Tables and Figures

1. Introduction

Noise-induced hearing loss is the second most common form of hearing deficit besides age related hearing loss, and can be observed even in younger people (Rabinowitz, 2000; Shargorodsky et al. 2010; Le Prell et al., 2011). Also hearing loss is a gradual process, which takes place over a long period of time, people are not always aware that it is taking place (Lipscom, 1972; Magrab, 1975; Le Prell et al., 2011; Ben-David et al., 2011). Losing the ability to hear clearly as a result of too much exposure to loud noise can be devastating and disabling, not least because noise-induced hearing loss could be prevented (Rabinowitz, 2000). All too often we take normal hearing for granted and cannot imagine how it would be to slowly lose the ability to hear clearly what is said around us. Generally, we hear sounds at a safe level that does not affect our hearing. It is only when we are exposed to loud noise that our hearing might suffer, because any loud noise can potentially cause hearing loss. One of the first signs of noise-induced hearing impairment is the reduced perception of higher frequencies which means that the ability to hear fricatives such as /s/, /f/ and /z/ is impaired (Magrab, 1975; Ben-David et al., 2011; Le Prell et al., 2011). High-frequency hearing loss also makes hearing in noisy environments more difficult, as sentence intelligibility decreases with increasing proportions of noise (Stilp and Kluender, 2010). Regrettably, we only start thinking about hearing loss when we lose the ability to hear well. This is extremely sad because hearing loss is an irreparable damage to the ear. Once the highly specialised ear cells are destroyed, “they do not regenerate and cannot be stimulated to regenerate; they are lost forever” (Magrab, 1975:35).

This paper proposes an acoustic phonetic based approach to the study of hearing loss by investigating how well normal hearing listeners are able to recognise fricatives in normal and simulated hearing-impaired conditions with and without background noise. The strategy is to present to listeners four voiceless fricatives (/f/, /θ/, /s/, /ʃ/) in the words *fin, thin, sin* and *shin* which are embedded in a carrier sentence and played to the listener in the four different conditions. The listeners’ results are recorded to analyse the impact of various conditions on the recognition ability of fricatives. Despite a rich history of research on consonant recognition in normal and noise condition (Miller and Nicely, 1955; Cole et al., 1996; Parikh and Loizou, 2005), in listeners with hearing loss (Skinner et al., 1997; Panda et al., 2010; Stilp and Kluender, 2010), in older and younger people (Kewley-Port et al., 2007; Nishi et al., 2010) and in normal hearing and simulated hearing-impaired conditions (Maniwa et al., 2008), the impact of various conditions tested simultaneously remains elusive.

1.1 Summary of the problem

Our hearing is exposed to noise emissions from different sources at most hours of the day, and some of these noises may have a detrimental effect on our hearing. Loud noises, in particular, can cause hearing loss either slowly or by constant exposure over a long period of time (Rabinowitz, 2000). The changes in hearing as a result of exposure to noise are complicated. “They include distortions of the clarity and quality of auditory experience as well as losses in the ability to detect sounds”(Magrab, 1975:32). Consequently, people may gradually lose the ability to hear sounds of high frequency which makes hearing or understanding of high-pitched voices such as women’s and children’s voices or fricatives particularly difficult (Ben-David et al., 2011; Le Prell et al., 2011).

Fricatives such as /f/, /θ/, /s/ and /ʃ/ are high-pitched sounds that are produced by rapid vibrations at high frequency. Vowels, by contrast, are low-pitched sounds that are produced by fewer vibrations, i.e. more slowly at lower frequency. Although in general high-pitched sounds are more difficult to distinguish, experiments have shown that fricatives can be recognised by their place of articulation and the acoustic energy used in their production. Hence /s/-/ʃ/ pairs are easier to identify than /f/-/θ/ pairs in both clear and conversational conditions, i.e. in a noise environment in both normal and simulated hearing conditions, as /s/ and /ʃ/ are louder than /f/ and /θ/ (Maniwa et al., 2008). These findings are congruent with Stevens et al. (1992), who showed that acoustic energy alone does not account for perceptual consonant recognition. It might be that /f/ and /θ/ are more frontal than /s/ and /ʃ/ and as such are recognised better (Dubno and Levitt, 1981). Collectively, these studies showed that listeners demonstrate difficulties understanding vowels and consonants in noise and impaired conditions.

1.2 Aim of the study

The present study investigates the recognition of four voiceless fricatives (/f/, /θ/, /s/, /ʃ/) in various conditions. The reason for focussing on fricatives only is that in comparison to other consonants, more energy is used in the production of fricatives which results in very high amplitude noise. This greater source strength contributes to the intelligibility enhancement and recognition of fricatives (Maniwa et al., 2008:123). Earlier findings, for example, have shown that /s/ can be distinguished from /ʃ/ by the energy used to produce the fricative, as the frequency range for /s/ (3.5 to 5 kHz) is higher than for /ʃ/ (2.5 to 3.5 kHz). Yet there seems to be some disagreement with regard to the recognition of /s/ in the hearing-impaired condition. According to Skinner et al. (1997), /s/ seems to be recognised correctly even in hearing-impaired conditions. Maniwa et al. (2008), on the other hand, noted that /f/-/θ/ pairs are recognised better than /s/-/ʃ/ pairs in the hearing-impaired condition. The present study therefore tries to find out which effects can be replicated in a different linguistic environment. Furthermore, as few research papers have looked at the relationship of all four variables under consideration (talker’s gender (2), noise (2), impairment (2) and fricatives (4)), the interaction between these four variables will be studied.

The study does not aim to present new methods in the field of consonant recognition but rather investigates whether the same results as in previous studies (Gordon-Salant, 1986; Zeng and Turner, 1990; Maniwa et al. 2007; Silbert and de Jong, 2008) can be replicated by testing students in a different language environment across several conditions. For this reason, analysis will merely focus on pitch, frequency range and frication duration for identification clues. The signals in the hearing-impaired conditions are filtered to imitate telephone speech. Hence, the following questions will be addressed: First, are there any significant differences to be observed in the overall recognition of fricatives? Second, will the overall results for the labio-dental and dental fricatives (/f/, /θ/) be better than for alveolar and palato-alveolar fricatives (/s/,/ʃ/) as suggested by Maniwa et al. (2008)? Third, do listeners also perform worse in the hearing-impaired condition (+noise, -noise) as various studies have shown? And finally, which fricatives have been confused most with each other in which conditions? Do the results comply with literature (Miller and Nicely (1955)?

1.3 Structure of the study

The paper is structured in such a way that the first part is dedicated to a brief introduction of fricatives and their distinguishing characteristics before the presentation of current research and literature in the field of noise-induced hearing loss and consonant recognition. The second part is supplemented by own personal research material which consists of recorded voice stimuli that have been prepared for use in E-prime. The results of the listening experiment are then statistically analysed using SPSS. For the conclusion in the third part, information from the first part and the research material are compared and discussed in detail. Finally, the last section deals with open questions and draws conclusions.

2. Literature

A full understanding of speech sounds is only possible when the incoming sounds can be processed adequately by the listener's auditory system (Hayward, 2000:130). In the case of a person suffering from high frequency hearing loss, this means that he or she has difficulty hearing and identifying consonants such as the fricatives /s/, /f/, /t/ and /z/ (Panda et al., 2010). For these people, it might be distressing not to be able to use pitch or duration contrastively anymore to distinguish among consonants. Hence, in the following the focus will be put on the recognition of fricatives in various sound conditions.

2.1 Fricatives

Fricatives are high-pitched noise-like sounds. They do not only constitute the largest set of consonants in the English sound system, they also, in comparison to other consonants, use more energy in their production which results in very high amplitude noise. Fricative consonants are distinguished from other speech sounds by their manner of production. This is to say in comparison to other consonants, more energy is used to produce fricatives. Fricatives are produced by forming a narrow constriction in the vocal tract. When air is blown through this constriction, it becomes turbulent in flow; and the acoustic result of this turbulence is noise (Stevens, 1997; Jongman et al., 2000). However, not all fricatives are produced the same way and with the same amount of air. In the case of voiceless fricatives, the vocal folds in the larynx are apart and air from the lungs is able to flow through freely. The turbulence in the airstream is only created in the frication with an obstacle such as the teeth, for example (Ladefoged, 2005). Accordingly, a sibilant is a type of fricative made by directing air through a narrow channel in the vocal tract towards the sharp edge of the teeth (such as /s/ and /ʃ/ in *sin* and *shin*), while non-sibilant fricatives (including /f/ or /θ/ as in *fin* and *thin*) produce their characteristic sound with the tongue or lips and the place of contact in the mouth, but without the involvement of the teeth (Ladefoged, 2005). In general, sibilants are louder than their non-sibilant counterparts, and most of their acoustic energy occurs at higher frequencies (Hayward, 2000). The fricatives /f/ and /θ/, for example, are characterised by a more even concentration of energy and shorter duration. /s/ and /ʃ/, on the other hand, have more and lower energy, but most of their acoustic energy occurs at higher frequencies making them louder and higher-pitched (Strevens, 1960; Brinton and Brinton, 2010). The voiceless alveolar fricative /s/, for example, is produced with on average 58.3 dB, whereas the overall sound pressure level for the voiceless labio-dental fricative /f/ amounts to around 53.0 dB (Badin, 1989:36). Hence the amount of frication and acoustic energy used in the production should help a listener identifying a consonant. Generally, /s/ has the most acoustic strength at around 6 to 8 kHz but can reach as high as 10 kHz. /ʃ/ tends to have most of its acoustic energy at around 3 to 4 kHz, but can extend up to around 7 to 8 kHz (Ladefoged, 2003). Yet the human auditory system is most sensitive to sounds at frequencies between 2 and 5 kHz (Johnson, 2004; Magrab, 1975).

2.2 Effects of noise on hearing

It is known that excessive exposure to loud noise leads to hearing loss and consequently can have a detrimental effect on a person's life (Magrab, 1975; Vogel et al., 2009). Lipscomb (1972) observed high frequency hearing impairment in the range of 2 - 6 kHz in students aged between 16 and 21 years. In his study, he tested 2769 new students at the University of Tennessee in 1968. They were given a modified screening test to assess their threshold for hearing. In 1969, the test was repeated with 1410 students. The findings from both tests were compared and revealed that "high frequency impairment rose dramatically" within the one year period (Lipscomb, 1972:231). In another study, Shargorodsky et al. (2010) tested young people aged 12 to 19 years by audiometric evaluation and discovered that 1 in 5 American adolescents demonstrated hearing loss. The authors also stated that high-frequency hearing loss was more common than low-frequency hearing loss, and that there was a significant increase to be observed between the time periods of 1988 to 1994 and 2005 to 2006. A more recent study by Le Prell et al. (2011) examined college students in North America. Their findings revealed that almost one-quarter of the students tested already suffered from hearing loss. In summary, these results raise concerns about the awareness of hearing loss in younger people because once a person's hearing is damaged; the impairment cannot be reversed but continues to deteriorate with age (Magrab 1975). This is particularly sad in view of the fact that the number of students suffering from hearing loss is steadily increasing, and that noise-induced hearing loss could be prevented by reducing the exposure to loud music (Vogel et al., 2009; Rabinowitz, 2000).

2.3 Impact of noise on understanding of speech

Exposure to noise does not only impair hearing but also impedes understanding of speech. A number of researchers have investigated vowel and consonant recognition, pointing out that noise impedes speech understanding (Magrab, 1975; Darwin, 2008; Ben-David et al., 2011). Accordingly, listeners encounter more difficulties understanding sentences uttered in a noisy environment than in a quiet environment as sentence intelligibility decreases with increasing proportion of noise (Stilp and Kluender, 2010). The impact of noise on consonant recognition has been studied by Parikh and Loizou (2005) who carried out an experiment in babble and speech-shaped noise conditions. The authors analysed stop consonants (b, d, g, p, t, k) only. Their results on consonants showed significant effects of noise but a non-significant effect of noise type. This suggests that any kind of noise can impair recognition of stop consonants. Despite the inference of noise, the authors found that stop-consonant identification still remained high (80% to 90% correct). The authors suspect that this is because the particular noises selected for the experiment mask the low frequencies more than the high frequencies

(Parikh and Loizou, 2005:3883). The relevance of these results for the present study is that plosives are characterised by sudden high frequency noise bursts. According to the energy used, the consonant might either be perceived as /b/ or /p/. In some speech context, this difference might lead to consonant confusion similar to fricatives. Different results have been reported by Nishi et al. (2010) who tested three groups of children between 4 and 9 years and adults (19-41 years old) on 15 English consonants embedded in vowel-consonant-vowel (VCV) nonsense syllables with the vowel /ɑ/. In their experiment, the carrier sentences were recorded by a male talker. The findings revealed that stops were the most problematic consonants and got confused most often by all subjects. Also, the fricatives /s/ and /ʃ/ got confused by all groups in almost all conditions. Another experiment on consonant confusion carried out by Wang and Bilger (1973) showed that all groups made place of articulation confusions in the presence of noise, in particular with /s/ and /ʃ/ and that overall performance for voiced consonants was relatively high for all groups compared to voiceless consonants.

As these studies have shown, listeners often show difficulty in distinguishing between consonants (Brinton and Brinton, 2010:31). Miller and Nicely (1955) have studied fricative confusions and found that in noisy speech, listeners mainly confuse /f/ with /θ/ or /ʃ/ with /s/. Table 1 below also shows that /s/ is confused with /f/, /θ/ or even /ʃ/, and /θ/ is confused with /f/ and /s/.

Table 1: Fricative confusions in noisy speech from Miller and Nicely (1955)

	/f/	/v/	/θ/	/ð/	/s/	/z/	/ʃ/	Total
/f/	199		46	1	4	-	-	250
/v/	-	177	1	29	-	4	-	217
/θ/	85	2	114		10	-	-	211
/ð/	-	64		105	-	18	-	187
/s/	5	-	38		170	-	10	223
/z/	-	4	-	22	-	132	-	158
/ʃ/	-	-	-		3	-	267	270

Results collected from 2,000 observations by Miller and Nicely (1955:341)

It has also been observed that the reason why /s/ is less often confused with /ʃ/ is because of the energy used to produce it. Furthermore, /s/ and /ʃ/ have louder frication noise than other fricatives and well-defined peaks (Johnson, 2004; Ladefoged, 2005). This means that the amount of acoustic energy can help distinguish an alveolar fricative (/s/) from a palato-alveolar fricative (/ʃ/), for example. Maniwa et al. (2008) showed that /s/, /z/, /ʃ/ and /ʒ/ are always easier to identify than /f/, /v/, /θ/, /ð/. Yet in the simulated hearing-impaired condition, /f/-/θ/ pairs were recognised better than /s/-/ʃ/ pairs.

In summary, these studies point out that it is difficult for a listener to understand and recognise vowels and consonants in noise conditions. Age differences in listeners have been investigated and the impact of excessive noise on hearing. In spite of the fact that we hear different voices and have to identify speech across different listening conditions, previous studies did not provide information regarding the impact of a talker’s gender on the recognition of fricatives in normal and noise conditions..

2.4 Simulated hearing impairment

People suffering from high frequency hearing loss often have a reduced dynamic range on speech recognition, meaning that they can merely hear sounds in a narrower frequency region than normal hearing people and in particular lose sensitivity of higher frequency sounds (Stuart et al., 1995). Hence simulating hearing loss is one method to learn about hearing impairment and to allow normal hearing listeners to perceive what it is like to be hearing-impaired. Furthermore, simulated hearing impairment allows for a controlled situation as it is often difficult to determine “which aspects of auditory processing contribute most to degraded speech reception” (Maniwa et al., 2008:1120). Hence, a common strategy is to simulate the effects of hearing impairment and have normal hearing listeners experience the impact of selected hearing loss (Maniwa et al., 2008). Also, simulated hearing impairment experiments with normal hearing listeners have delivered findings that are comparable to subjects with cochlear impairment who typically suffer from high frequency loss. The performance of simulated hearing-impaired listeners in noise condition diminishes relative to the normal hearing people (Stuart et al., 1995). Skinner et al. (1997) explored speech recognition abilities of ten adults with hearing impairments. The authors presented listening material to the candidates at different sound-pressure levels and found that the fricative /s/ is identified correctly in 60% of all cases even at the lowest sound level. Hence, the findings show that there is a tendency for louder and higher-pitched fricatives to be recognised well even in hearing-impaired conditions.

2.6 Talker’s voice

Female voices are on average higher pitched than male voices (Ben-David et al., 2011; Le Prell et al., 2011). The reason for it being that male focal folds “tend to be longer and thicker than female vocal folds causing them to vibrate more slowly” and hence produce lower pitched frequencies (Simpson, 2009:622). Generally, the fundamental frequency (F0) for the male voice lies between 100 Hz and 150 Hz. Female vocal folds are shorter and lighter and vibrate at approximately twice the male frequency; this is to say between 200 and 220 Hz. Many papers have investigated gender identification by listeners (Fu et al., 2004; Sheffert, 1998; Hillenbrand and Clark, 2009). But there is scarcely any literature available regarding the effect of voice gender on consonant recognition. The study by Mackersie et al. (2011) demonstrated that vocal tract length has an impact on sentence comprehension. Performance improved for people with normal hearing and hearing loss respectively, but only when the target speaker had the higher fundamental frequency (F0). In other words, listeners’ performance improved when the talker had a shorter vocal tract length, which is typical for women. Ferguson

(2004) investigated vowel intelligibility for normal-hearing listeners in normal and conversational speech when exposed to different speakers. Her findings showed that speech modification can also influence a listener's intelligibility of sentences. This is to say that sentences produced by speakers who, for example, increase their vowel duration or speak slower are perceived better in clear speech. These findings suggest that not only a speaker's frequency range but also the way utterances are produced influence a listener's speech perception. Hence, a number of acoustic features help listeners to discriminate voice gender (Fu et al., 2004; Ferguson, 2004; Friedrich et al., 2008; Simpson, 2009).

In summary, consonants can be distinguished by their place of articulation and their acoustic energy. Hence, the fricatives /f/ and /θ/ tend to be less loud than /s/ or /ʃ/, have a more even concentration of energy and shorter frication duration. In the production of /s/ and /ʃ/, on the other hand, more and lower energy is used but at higher frequencies, which makes them louder and higher-pitched. Research has further shown that some information on fricatives is retained even under various noise conditions, which suggests that listeners should be able to identify and recognise fricatives according to their particular acoustic features (Maniwa et al. 2007). Yet if hearing is impaired, not all of these acoustic clues are available to listeners anymore. Research revealed that high frequency hearing loss as a result of exposure to loud music disrupts spoken language comprehension (Ingram, 2007:53). Moreover, background noise might cause additional speech understanding problems.

3. Method

3.1 Listeners

Thirteen female students (age: *M* = 24.5 years, *SD* = 3.76) and thirteen male students (age: *M* = 25 years, *SD* = 3.85) from the University of Edinburgh participated in the experiment. The listeners were all native speakers of British English residing in Scotland (female: 2 Irish, 3 Scottish, 8 English; male 5 Scottish, 8 English). They reported normal hearing and no history of speech or language disorders. The listeners were paid for participating in the experiment.

3.2 Stimuli

Four English voiceless fricatives /f/, /θ/, /s/ and /ʃ/ were used to form monosyllabic words with the fricative in initial position followed by the close high-front vowel /i/ and ending in the voiced alveolar nasal /n/ to form *fin, thin, sin* and *shin*. The fricatives were in word initial position to eliminate the possibility of co-articulation from preceding vowels which may provide “enough information about adjacent consonants to allow listeners to recover the intended words” (Cole et al., 1996:855). Moreover, studies have shown that recognition of consonants in initial positions is significantly higher than recognition of syllable final consonants (Dubno and Levitt 1981; Cheung et al., 2001). Furthermore, real words instead of nonwords have been selected because research has shown that listeners recognised words better than nonwords (Rubin et al., 1976:394).

The stimuli to be recorded were embedded in the carrier sentence *I say* ______ *now*. Four repetitions of each carrier sentence were recorded twice resulting in a total of 32 sentences per talker. The order of the stimuli in the carrier sentence alternated [(thin, fin, shin, sin); (sin, shin, thin, fin); (fin, sin, shin, thin); (shin, thin, sin, fin)] to account for any tiring or repetition effect during the recording, which might have influenced the voice quality of the talker. For the experiment, only one carrier sentence was selected (4 sentences per person).

3.3 Recording of stimuli

The recording took place in a single-wall, sound-attenuating booth where the talker was seated approximately 20 cm from the Shure SM7b dynamic cardoid pattern microphone. The microphone output was routed to a preamplifier (Alice Mic Amp Pak) and then to an iMac running ProTools using a Digidesing 003 audio interface. All recordings were sampled at 48 kHz (16 bit rate) and saved as a wav (mono) file. The gain was adjusted manually in order to match intensity differences among speakers. The duration of the carrier sentence was fixed at 2000 ms. The sample rate was reduced to 22.05 kHz for use in E-prime version 2.0.8.90.

3.4 Voice recordings

The talkers for the voice recording were selected for voice quality, regional accent, age and smoking history. The sentences were recorded by 1 female native Scottish English speaker from Edinburgh (age 29, former smoker) and 1 male native Scottish English speaker from Edinburgh (age 31, former smoker) both recruited from the University of Edinburgh. The voice recordings have been analysed in PRAAT 1.2.3 using the start and stop point of each consonant (f/, /θ/, /s/, /ʃ/) to measure the duration and average pitch. The duration of the frication noise in the four fricatives (f/, /θ/, /s/, /ʃ/) recorded in the carrier sentence by a female talker and a male talker are illustrated in Table 2 below.

Table 2: Duration of frication noise

	Sibilants		**Non sibilants**	
	/s/	/ʃ/	/f/	/θ/
Female voice	0,117 seconds	0,117 seconds	0,102 seconds	0,115 seconds
Male voice	0,147 seconds	0,140 seconds	0,145 seconds	0,148 seconds

In both the male and the female voice, the fricatives /s/ and /θ/ have similar duration of frication noise. On average, the duration of the frication noise in the female voice is longer for sibilants than for non sibilants. In the male recording, the duration of the frication noise for non sibilants is longer than for sibilants. The pattern of the recorded male voice does not conform to literature, which argues that /f/ and /θ/ are characterised by a more even concentration of energy and shorter duration than /s/ and /ʃ/ (Pirello et al., 1997).

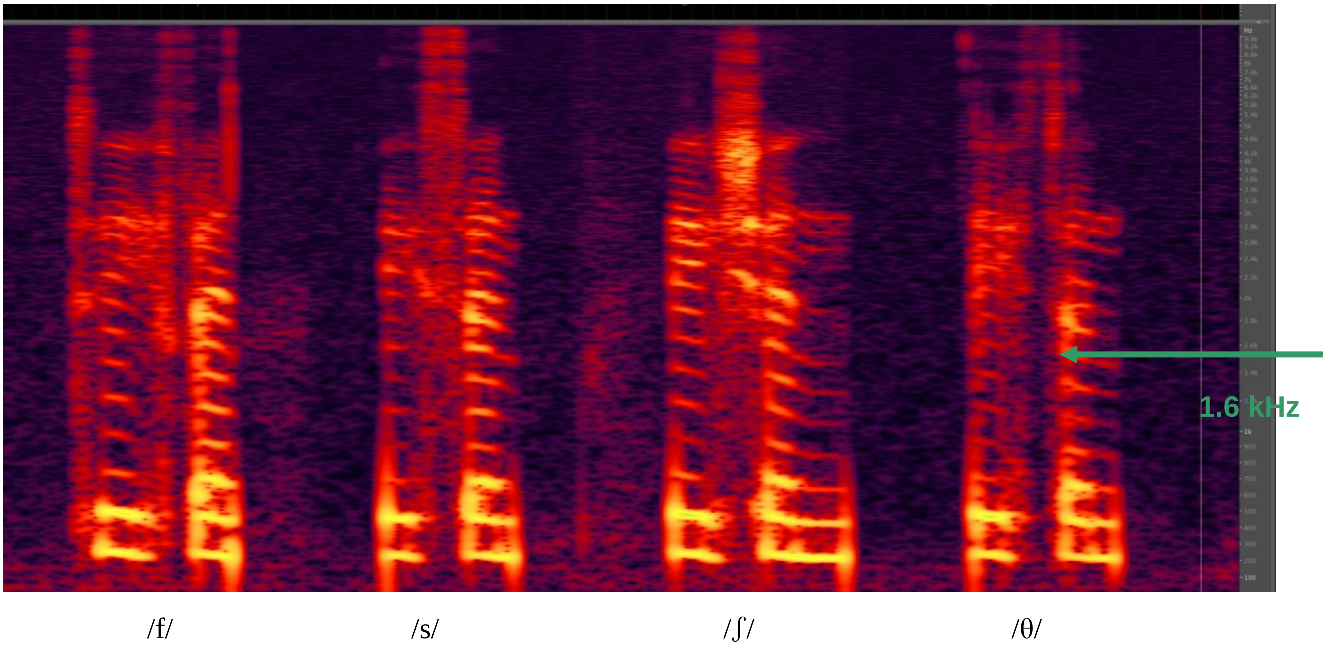
Table 3: Average acoustic pitch

	Sibilants		**Non sibilants**	
	/s/	/ʃ/	/f/	/θ/
Female voice	68.76 dB	74.74 dB	69.09 dB	66.28 dB
Male voice	73.46 dB	75.75 dB	61.84 dB	64.86 dB

Regarding pitch measurements, the values for /s/ and /ʃ/ are higher than those for /f/ and /θ/ in both the female and the male recordings (see Table 3). These results comply with literature stating that the sibilants /s/ and /ʃ/ tend to be much stronger than the rest of the fricatives (Ladefoged, 2003). The following table shows the average acoustic energy used in the production of the recorded fricatives.

The frication range of the male voice recordings lies between 100 Hz and 4.6 kHz and the female voice recordings between 100 Hz and 5.4 kHz. Both, male and female voices reach peaks above 10 kHz when producing the fricatives f/, /θ/, /s/, /ʃ/ in the words *fin*, *thin*, *sin* and *shin.* Figure 1 below shows the spectrogram of the female voice recordings of the four words embedded in the carrier sentence “I say_____now” recorded by the female talker.

Figure 1: Spectrogram female voice recordings (fin, sin, shin, thin)

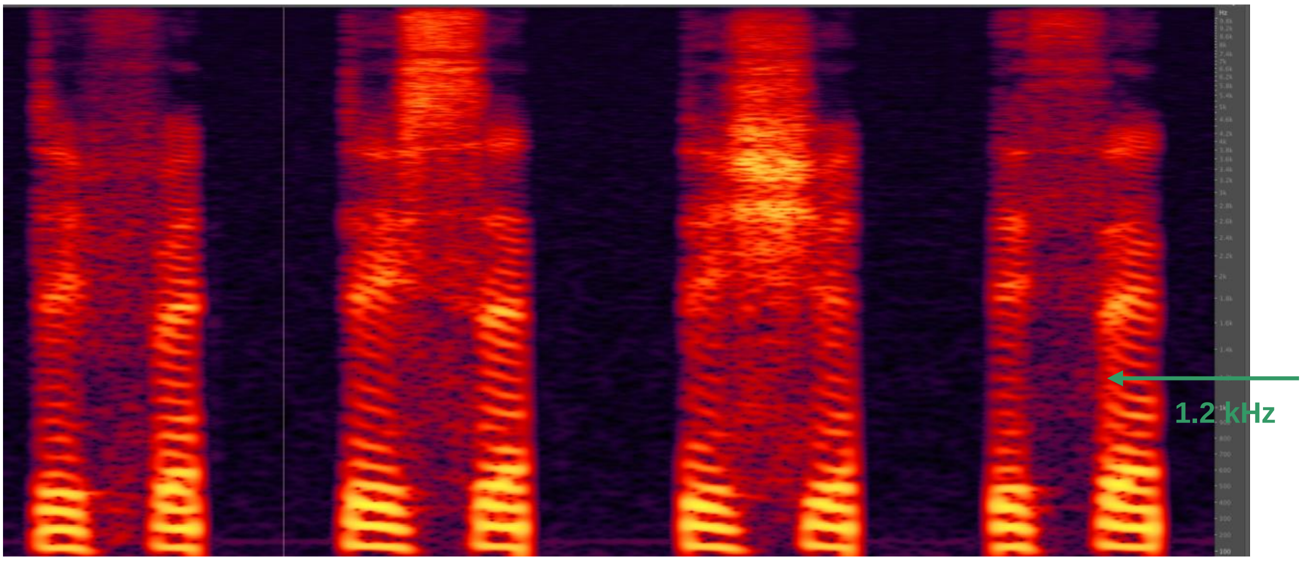


The spectrograms have been generated in Audition version 2. They show the frequency content as a function of time. Frequency is on the vertical axis and time is on the horizontal. The orange colour is an indicator of energy used in the production of the speech sample, hence the darker it gets, and the less energy has been used. The spectrogram also shows that /θ/ has the least energy and /f/ only a little bit more; whereas /s/ and /ʃ/ have a lot of energy.

The spectrogram further illustrates that /f/ has a rather even distribution of energy above 1.6 kHz, but also shows some energy below 1.6 kHz. The fricative /θ/ also has a fairly uniform distribution of energy above 1.6 kHz, but it does not show energy below 1.6 Hz. The fricative /s/, on the other hand, shows greater concentration of energy above 2 kHz; and /ʃ/ has its highest energy concentration between 2 kHz and 4 kHz.

Figure 2 shows the spectrogram of the male voice recording. Each of the four words (fin, thin, sin and shin) has been embedded in the carrier sentence “I say_____now”.

Figure 2: Spectrogram male voice recordings (fin, sin, shin, thin)



/f/ /s/ /ʃ/ /θ/

The spectrogram illustrates that /f/ has a rather even distribution of energy above 1.2 kHz, but also shows some energy below 1.2 kHz. The fricative /θ/ also has a fairly uniform distribution of energy above 1.2 kHz, but it does not show energy below 1.2 kHz. The fricative /s/, on the other hand, shows greater concentration of energy above 4 kHz; and /ʃ/ has its highest energy concentration between 1.8 kHz and 5 kHz.

According to literature, the voiceless alveolar fricative /s/ tends to show a primary spectral peak at around 4 to 5 kHz with most of its acoustic strength at around 6 to 8 kHz, but can reach as high as 10 kHz. In the production of the voiceless palato-alveolar fricative /ʃ/, on the other hand, the acoustic energy is concentrated at around 2 to 4 kHz and can reach up to around 8 kHz (Stevens, 1997:500, Ladefoged, 2003). The fricative /f/ tends to have a flatter spectrum than other the other fricatives /s/, /θ/ and /ʃ/. The voice recordings used in the present study show similar spectral characteristics.

3.5 Design

There are two groups to be tested (male and female listeners). The experiment uses a within group design, with voice sex (male vs. female), acoustic interference (noise vs. no noise), hearing impairment (normal hearing vs. impaired hearing) and the word (thin, fin, sin or shin) as independent variables, and recognition accuracy (RA) and reaction time (RT) as the dependent variables.
The 2 x 2 x 2 x 4 [Voice (male, female), Condition (normal, impaired), Noise (+noise, -noise), Target word (fin, thin, sin, shin) design produced 32 factors. The experiment was divided in two blocks, the experiment with 32 sentences and the repetition session with the same 32 sentences each time presented in random. The carrier sentences were presented randomly in such a way that no target word

is presented twice in a sequence to the listeners. Also, the target words displayed on the computer screen were presented in a different sequence every time they heard a new sentence. The programming and running of the experiment was done in E-Prime version 2.0.8.90.

3.6 Background noise

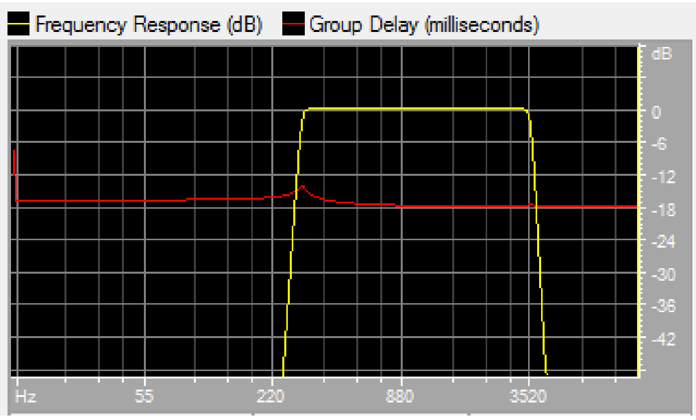
There are no pure signals. There are always interferences. Noise is a summation of unwanted or disturbing energy from natural and sometimes man-made sources. It is often generated deliberately to be used in experiments as a test signal. In the present study, the noise condition was simulated in E-Prime by using 2000 ms of pink noise ($1/f$ noise) to cover the duration of the carrier sentence recorded. Pink noise has a power spectral density that falls at 3dB/octave with rising frequency and is therefore more useful in audio testing because it contains constant energy per octave (Magrab, 1975). The following settings have been used:

Normal condition:	Noise attenuated by 100dB, i.e. effectively muted, meaning 100% signal, 0% noise
Normal + noise condition:	50% signal, 50% noise
Impaired condition:	Noise attenuated by 100db, i.e. effectively muted, meaning 100% signal, 0% noise
Impaired + noise condition:	50% signal, 50% noise

3.7 Filtered signal

Simulating hearing loss is one method to learn about hearing impairment and to allow normal hearing listeners to perceive what it is like to be hearing-impaired. Therefore, the impairment filter was set to cut off frequencies below 300 Hz and above 3.5 kHz. The audio files were normalised to give them the same maximal level and the same volume using the normalised function in Audion version 2. Consequently, this band-pass filter simulates a high-frequency hearing-impaired condition similar to a telephone bandwidth (300 to 3 kHz). Figure 3 illustrates the band-pass filter used in the experiment.

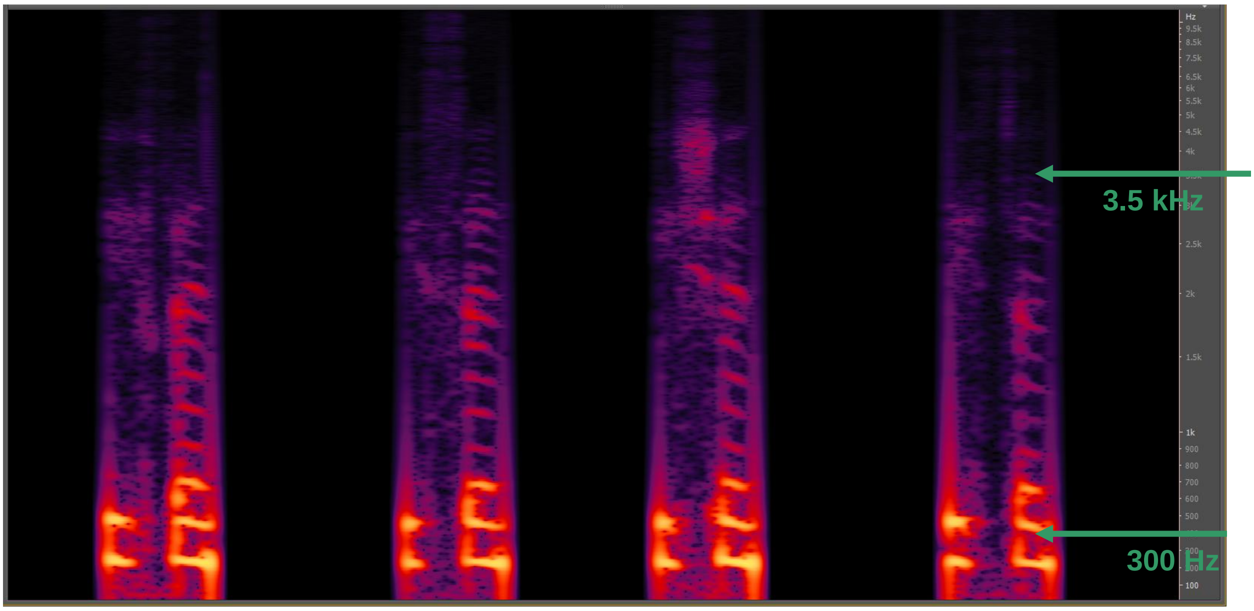
Figure 3: Scientific filter setting



A band-pass filter is like a combination of a low-pass filter and a high-pass filter and allows for the modulation of some aspects of speech and hearing (Johnson, 1997:19).

Figure 4 shows the spectrogram of the female voice recordings with the band-pass filter in the "impaired" condition. The four fricatives are embedded in the carrier sentence "I say _____ now". Compared to the undistorted recordings – see Figure 1 – the signals are weaker and most of the energy is concentrated in the lower frequency regions.

Figure 4: Spectrogram female voice recordings, filtered condition

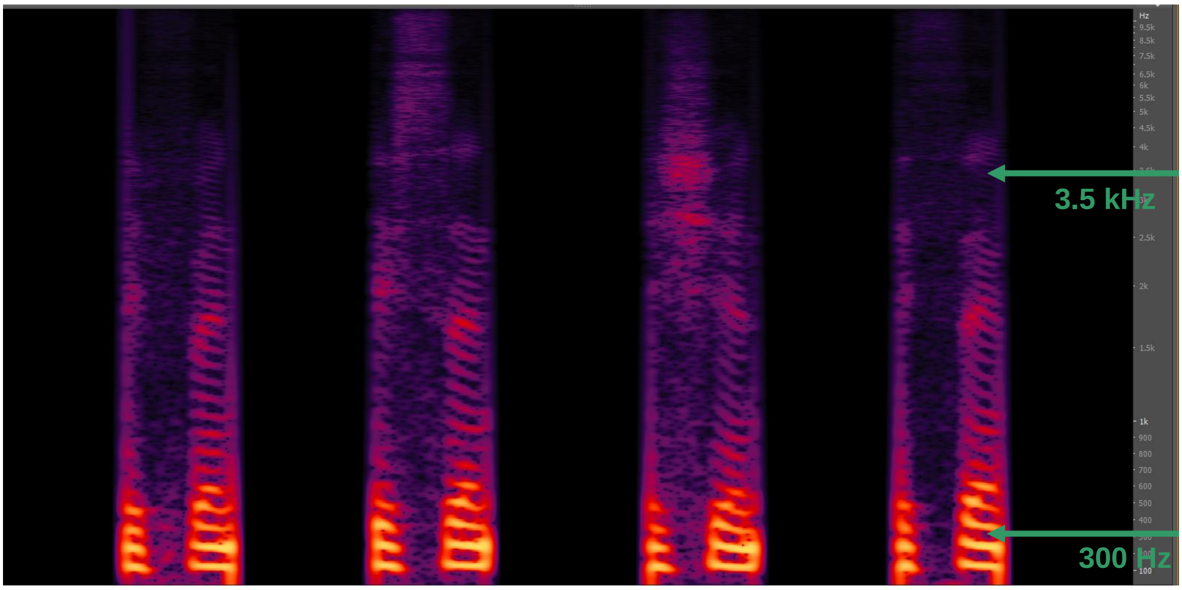


/f/ /s/ /ʃ/ /θ/

The spectrogram shows frequency content as a function of time. Frequency is on the vertical axis and time is on the horizontal.

Figure 5 below shows the band-pass filtered spectrogram of the male recordings; this is to say in the "impaired" condition. The four fricatives are embedded in the carrier sentence "I say _____ now" and recorded by the male talker.

Figure 5: Spectrogram male voice recordings, filtered condition



/f/ /s/ /ʃ/ /θ/

3.8 Procedure

Listeners were tested individually in a single-wall, sound-attenuating booth and informed that they would have to listen to carrier sentences with the target word embedded in the middle in various conditions and make a decision on the target word they think they have heard. The listeners were invited to listen to the instructions given by the investigator before starting with the experiment. The experiment was divided into two parts with a practice session proceeding the experimental sessions. In both sessions of the experiment the task for the listeners was the same: to identify the consonant in initial position as one of the fricatives f/, /θ/, /s/, /ʃ/ in a four-alternative forced-choice identification task (Maniwa et al. 2008).

Testing was conducted in a single-wall, sound-attenuating booth on a Toshiba Satellite Pro Laptop with a 20-inch liquid crystal display. The recordings were played to the listener over a set of Sennheiser eH2270 headphones with an output set to 70 dB on average. The entire experiment was completed in one 10 minute session.

Each stimulus was presented in the carrier sentence *I say* ______ *now* and played in random order to the listener. At the same time, the four stimuli used in the experiment (thin, fin, sin, and shin) were displayed on the computer screen. The order of the stimuli on the screen was constantly alternated so as not to tempt the listeners to always choose the same stimulus. A practice condition was conducted prior to the experiment in to familiarise the subjects with the stimuli and the experiment. In the practice session, all four randomised baseline stimuli were presented in neutral condition (normal-noise) to the listener. No feedback was provided to the listeners. The immediately following experiment was divided into two parts. First, the 32 carrier sentences were presented in all four possible combinations (2 voices, 2 noise conditions, 2 hearing conditions). Second, the experimental condition was immediately preceded by a repetition session where all 32 carrier sentences were presented again in random order and all possible conditions. In total, the listeners heard 32 stimuli twice resulting in a total of 64 carrier sentences.

4. Findings

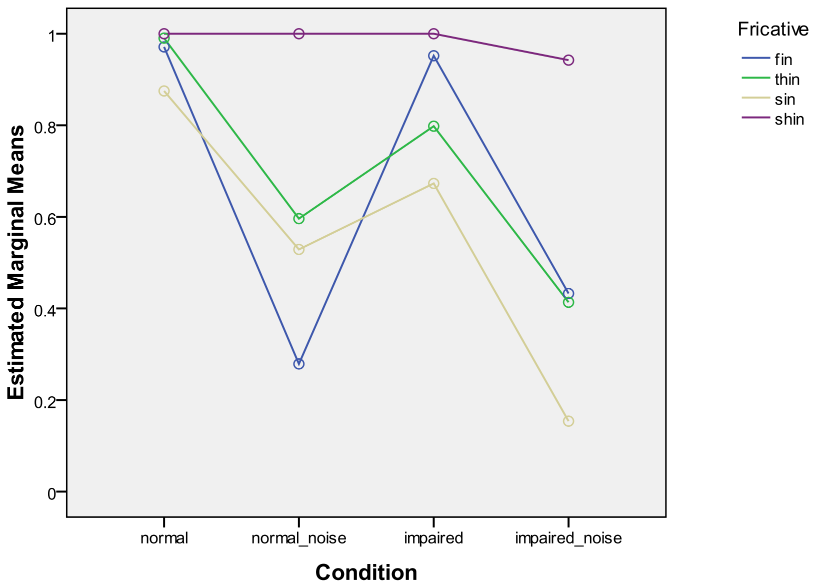
Each fricative was played twice to the listener; this is to say once in the male voice and once in the female voice in every condition, which amounts to 128 stimuli in total. For the present study, 26 subjects took part in the experimental and the repetition sessions and heard 256 stimuli in total.

A 2 x 2 x 2 x 4 repeated measures ANOVA was performed with 'Voice' (female, male) (2), 'Noise' (+noise, -noise) (2), 'Impairment' (+compressed signal, -compressed signal) (2), 'Fricative' (/f/, /θ/, /s/, /ʃ/) (4) as within subject factors and 'Results' (correct, incorrect) (2) as between subject factor. The result of Mauchly Tests for Sphericity shows $p > 0.05$ for within subject factors 'Voice and Noise' [$F(1,25) = 15.799$, $p < .001$], 'Voice * Fricative [$F(3, 75) = 33.653$, $p < .001$], 'Noise * Fricative' [$F(3, 61.259) = 19.366$, $p < .001$] and 'Noise * Impairment * Fricative' [$F(3, 75) = 6.091$, $p < .001$] revealed a main effect on the performance of listeners. Yet Mauchly's test also indicated that the assumption of sphericity had been violated in some cases, meaning that each participant is affected entirely differently by the manipulation of certain conditions. Therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity. The factors mainly affected are 'Voice * Impairment' [$F(1, 25) = .308$, $p = .584$, $\hat{\varepsilon} = 1.000$] and 'Noise * Impairment' [$F(1, 25) = .121$, $p = .730$, $\hat{\varepsilon} = 1.000$].

4.1 Effect of conditions

A Univariate Analysis of Variance with 'Condition' (4) and 'Fricative' (4) revealed a highly significant effect on 'Results' (2) [$F(9,1648) = 25.302$, $p > 0.05$). Figure 6 below illustrates the results and shows a tendency for /ʃ/ to be recognised best across all conditions.

Figure 6: Recognition of fricatives in various conditions



To find out how the individual conditions affected the recognition of fricatives, a frequency test including Pearson Chi-Square was carried out. The results showed a significant effect of ‘Condition’ on ‘Fricative’. In detail, the ‘normal’ condition [χ^2 (3, N = 416) = 56.246, $p < .001$] and the ‘impaired’ condition [χ^2 (3, N = 416) = 26.188, $p < .001$] least affected the recognition of fricatives. Most of the listeners made mistakes in the ‘normal+noise’ condition [χ^2 (3, N = 416) = 116.320, $p < .001$] and in the ‘impaired+noise’ condition [χ^2 (3, N = 416) = 135.998, $p < .001$]. The significance level also indicates that the difference could not simply have happened by chance. It seems more likely that the auditory conditions influence the recognition of fricatives.

The test design was set out as a forced-choice experiment. Hence, errors could only be committed by selecting the incorrect fricative. Tables 4 to 7 below illustrate how often and in which condition listeners recognised either correctly or incorrectly each of the fricatives.

The findings for the fricative /f/ in *fin* in Table 4 show that it has most often been confused with /θ/ in the normal+noise condition and also in the impaired+noise condition. The results further show that some listeners did not even recognise /f/ in the normal condition. Posthoc analysis using Tukey HSD revealed that the results in the normal+noise condition significantly differ from the results in the impaired+noise condition ($p = .013$).

Table 4: Fricative = FIN (Answer / Result)

				Answer				
Condition	Result			fin	thin	sin	shin	Total
Normal	Incorrect	/f/ in *fin*	Selected		2	1		3
	Correct	/f/ in *fin*	Selected	101				101
Normal+noise	Incorrect	/f/ in *fin*	Selected		64	9	2	75
	Correct	/f/ in *fin*	Selected	29				29
Impaired	Incorrect	/f/ in *fin*	Selected		5			5
	Correct	/f/ in *fin*	Selected	99				99
Impaired+noise	Incorrect	/f/ in *fin*	Selected		34	24	1	59
	Correct	/f/ in *fin*	Selected	45				45
TOTAL								416

In Table 5 below, the results for the fricative /θ/ in *thin* are presented and illustrate that noise considerably affected the recognition of /θ/, as it had most often been confused in the impaired+noise condition and in the normal+noise condition.

Table 5: Fricative = THIN (Answer / Result)

Condition	Result			Answer: fin	thin	sin	shin	Total
Normal	Incorrect	/θ/ in *thin*	Selected	1				1
	Correct	/θ/ in *thin*	Selected		103			103
Normal+noise	Incorrect	/θ/ in *thin*	Selected	29		8	5	42
	Correct	/θ/ in *thin*	Selected		62			62
Impaired	Incorrect	/θ/ in *thin*	Selected	13		8		21
	Correct	/θ/ in *thin*	Selected		83			83
Impaired+noise	Incorrect	/θ/ in *thin*	Selected	35		21	5	61
	Correct	/θ/ in *thin*	Selected		43			43

Posthoc analysis using Tukey HSD showed that the results in the normal+noise condition do significantly differ from the results in the impaired+noise condition (p = .007).

The results for the fricative /s/ are presented in Table 6. The figures show that /s/ has not been recognised well in the impaired+noise condition.

Table 6: Fricative = SIN (Answer / Result)

Condition	Result			Answer: fin	thin	sin	shin	Total
Normal	Incorrect	/s/ in *sin*	Selected	1	11		1	13
	Correct	/s/ in *sin*	Selected			91		91
Normal+noise	Incorrect	/s/ in *sin*	Selected	19	26		4	49
	Correct	/s/ in *sin*	Selected			55		55
Impaired	Incorrect	/s/ in *sin*	Selected	5	26		3	34
	Correct	/s/ in *sin*	Selected			70		70
Impaired+noise	Incorrect	/s/ in *sin*	Selected	26	45		17	88
	Correct	/s/ in *sin*	Selected			16		16

Posthoc analysis using Tukey HSD revealed that the results in the normal+noise condition differ considerably from the results in the impaired+noise condition (p < .005). Moreover /s/ has also been identified incorrectly in the normal condition more often than any of the other fricatives.

Table 7: Fricative = SHIN (Answer / Result)

Condition	Result			Answer: fin	thin	sin	shin	Total
Normal	Correct	/ʃ/ in *shin*	Selected				104	104
Normal+noise	Correct	/ʃ/ in *shin*	Selected				104	104
Impaired	Correct	/ʃ/ in *shin*	Selected				104	104
Impaired+noise	Incorrect	/ʃ/ in *shin*	Selected	2	3	1		6
	Correct	/ʃ/ in *shin*	Selected				98	98

The figures for the fricative /ʃ/ in Table 7 show that it had always been identified correctly apart from the impaired+noise condition where it had been confused with other fricatives.

The findings revealed that overall, the fricative /ʃ/ in the word *shin* was recognised best (98.6%), followed by /θ/ in the word *thin* (70%) and /f/ in the word *fin* (65.9%). The fricative /s/ in word *sin* received the lowest score of correct answers; this is to say 55.8%.

4.2 Consonant confusion

Table 8 illustrates the number of fricatives played to all listeners (32 times) and the number of times each of the fricatives had been recognised correctly or confused with another. The results differ considerably across the fricative [(fin/thin 25%), (fin/sin 8%), (fin/shin 0.7%)], [(thin/fin 12%), (thin/sin 9%), (thin/shin 2%)], [(sin/fin 12%), (sin/thin 26%), (sin/shin 6%)], [(shin/fin 0.5%), (shin/thin 0.7%), (shin/sin 0.2%)].

Table 8: Consonant confusion

		Listeners' answer: fin	thin	sin	shin	Total
Fricative	fin	274	105	34	3	416
	thin	78	291	37	10	416
	sin	51	108	232	25	416
	shin	2	3	1	410	416
Total		405	507	304	448	1664

Posthoc analyses using Tukey HSD indicated that the fricatives /f/ and /θ/ did indeed get confused rather often (p =.492) with each other. There is also a significant effect of /f/ on /s/ and /ʃ/ (p <.005).

The fricative /θ/ got rather often confused with /f/ (p = .492) but seldom with /s/ or /ʃ/ (p <.005). The results, however, also show no significant consonant confusion for /s/ with /ʃ/ (p < .005).

4.3 What are significant confusions?

4.3.1 Effect of voice

A Frequency test by subject showed that in the male voice condition, 77.9% of the fricatives were identified correctly, whereas in the female voice condition only 69.0%. The difference proved to be significant. A 2x4 ANOVA with 'Voice' (male, female) and 'Fricative' (f/, /θ/, /s/, /ʃ/) was carried out to analyse how accurate listeners were in recognising fricatives in the two voice conditions. The results revealed a significant effect of 'Voice' on fricative recognition [$F(3,1664)= 29.288$, $p<.005$].

Figure 7 below illustrates the recognition of the four fricatives in general across all the conditions when played in the male voice condition and the female voice condition.

Figure 7: Recognition of fricatives in male and female voice condition

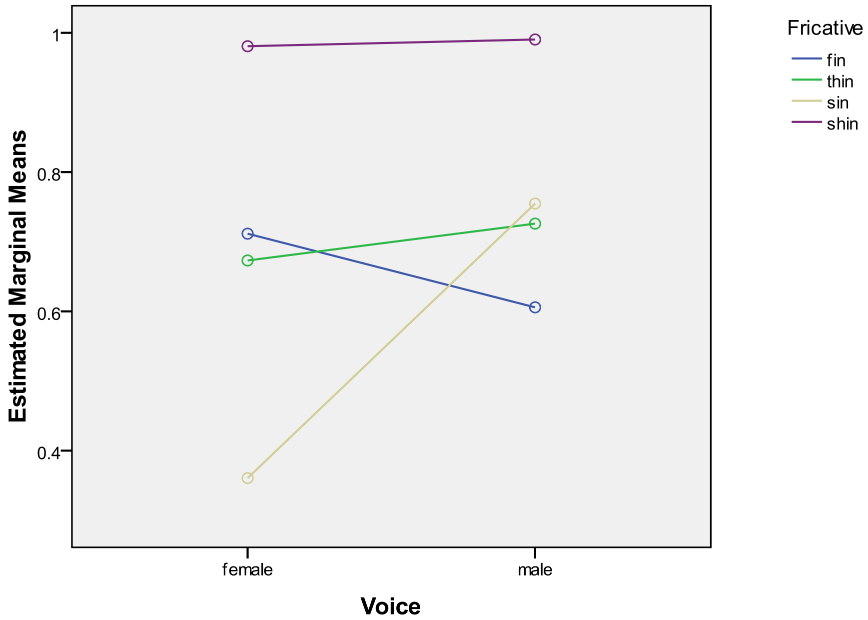


Figure 7 also shows a tendency for fricatives to be identified better in the male voice condition than in the female voice condition, with the exception of /θ/. In addition, it is clearly visible that /s/ in *sin* has not been recognised well in the female voice condition. /ʃ/, on the other hand, had been recognised well in both the male and the female voice condition.

A comparison of means confirmed that the fricative /ʃ/ in *shin* (*M*=.99, SD=.119) had been identified most accurately in both the male and female voice condition. The fricative /f/ in *fin* (*M*=.66, SD=.475) and the fricative /θ/ in *thin* (*M*=.70, SD=.459) received almost equal scores of correct recognition. There is, however, a considerable difference to be observed as regards the fricative /s/ in *sin* (*M*=.56, SD=.497), which received the lowest overall score of correct recognition.

The impact of the talker's voice on the recognition of fricatives is also statistically significant. A 2 x 2 One-Way ANOVA was carried out to test the statistical significance of 'Voice' / 'Results'. The findings showed a highly significant effect of 'Voice' on 'Results' [$F(1,1663) = 16.213$, $p < 0.005$], meaning that the talker's voice does indeed highly influence the recognition of fricatives. A more detailed analysis split by voice gender shows a statistically significant difference in the listener's performance. Fricatives have been identified much better in the male voice condition [$F\ (1,831) = 8.501$, $p = 0.004$] than in the female voice condition [$\mathrm{F}(1,831)= 7.719$, $p = 0.006$].

4.3.2 Effect of condition

A 2 x 2 x 2 x 4 repeated measures ANOVA was performed with 'Voice' (2), 'Noise' (2), 'Impairment' (2), 'Fricative' (4) between subject factors and 'Results'(2) and revealed that the 'Voice' of the talker as well as the 'Noise' condition do not seem to have a significant effect on the listener's performance in the 'Impairment' condition. Furthermore, the analyses did not reveal a significant main effect of 'Voice*Impairment' $F(1,25)= .308$, $p = .584$, MSE .015 (Huynh-Feldt correction); 'Noise*Impairment' $F(1,25)= .121$, $p = .730$, MSE .008 or 'Voice*Noise*Impairment' $F(1,25)= 7.046$, $p = .014$, MSE .782.

Figure 8: Fricative recognition 'No noise' condition

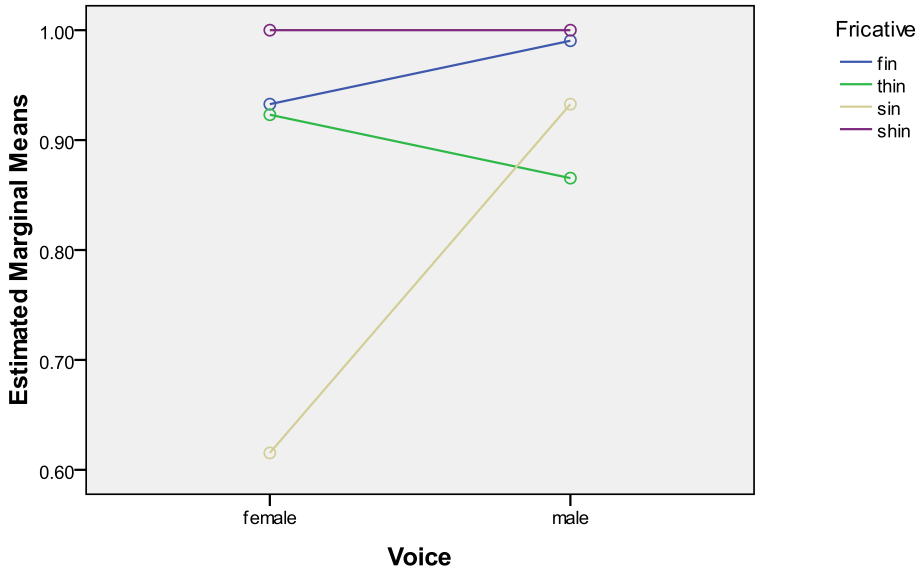


Figure 8 illustrates the recognition of fricatives in the ‘No noise’ condition (normal and impaired combined). The line graph shows that the recognition of fricatives was higher in the male voice condition than in the female voice condition. A Univariate Analysis of Variance with ‘Voice’ (2) and ‘Fricative’ (4) further revealed that in the ‘No noise’ condition, the influence of ‘Voice’ on the recognition is highly significant [$F(3,824) = 20.178$, $p < .005$].

Figure 9: Results / Fricative (no noise)

Voice			Fricative: fin	thin	sin	shin	Total
female	Results	incorrect	7	8	40	0	55
		correct	97	96	64	104	361
	Total		104	104	104	104	416
male	Results	incorrect	1	14	7	0	22
		correct	103	90	97	104	394
	Total		104	104	104	104	416

In the ‘No noise’ condition (normal, impaired), listeners overall identified 90.7% of all the fricatives correctly – see Figure 9 for illustration. In detail, 86.8% of the fricatives were recognised accurately in the female voice condition and 94.4% in the male voice condition. To find out how the talkers’ voices affected the recognition of fricatives, a Pearson Chi-Square test was carried out. The results showed that in the ‘No noise’ condition, listeners had more difficulties recognising fricatives in the female voice condition [χ^2 (3, N = 416) = 80.183, $p < .001$] than in the male voice condition [χ^2 (3, N = 416) = 27.410, $p < .001$]. The significance level also indicates that the difference could not simply have happened by chance. It is more likely that the female voice did indeed have an impact on the recognition of fricatives.

Figure 9: Fricative recognition ‘Noise’ condition

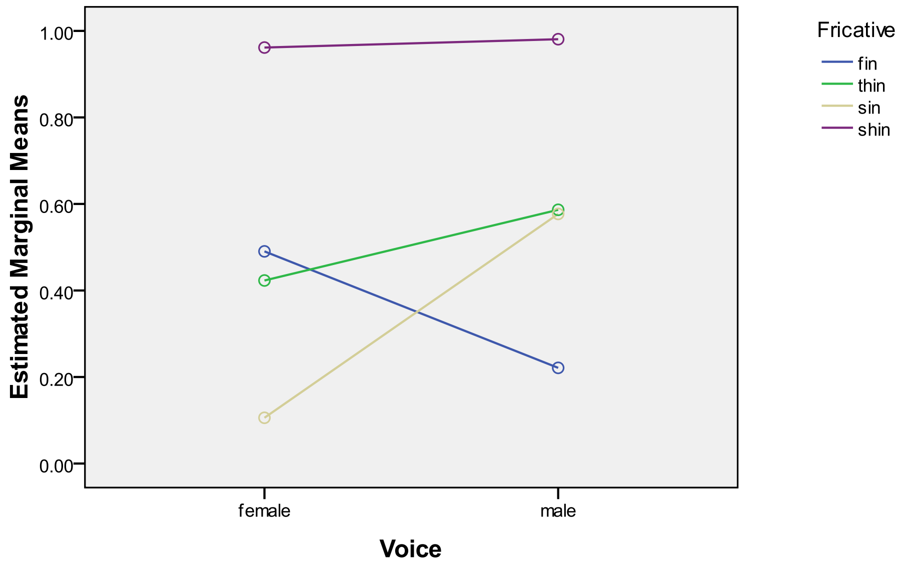


Figure 9 illustrates the recognition of fricatives in the 'Noise' condition when analysed by talker. The line graph shows the apart from /f/, all fricatives have been recognised better in the male voice condition. A Univariate Analysis of Variance with 'Voice' (2) and 'Fricative' (4) revealed that in the 'Noise' condition, the influence of 'Voice' on the recognition of fricatives is even more significant [$F(3,824) = 29.988$, $p < .005$] than in the 'No noise' condition.

Also, a Pearson Chi-Square test was carried out to see how each of the voices influenced the recognition of fricatives in the 'Noise' condition. The results showed that the female voice [χ^2 (3, N = 416) = 155.745, $p < .001$] had a far greater impact on the recognition of fricatives than the male voice [χ^2 (3, N = 416) = 124.342, $p < .001$] in the 'Noise' condition. In summary, the difference between the voices is considerably bigger in the 'No noise' condition.

Figure 10 presents an overview of how the four fricatives have been recognised in all the conditions separately.

Figure 10: Overview

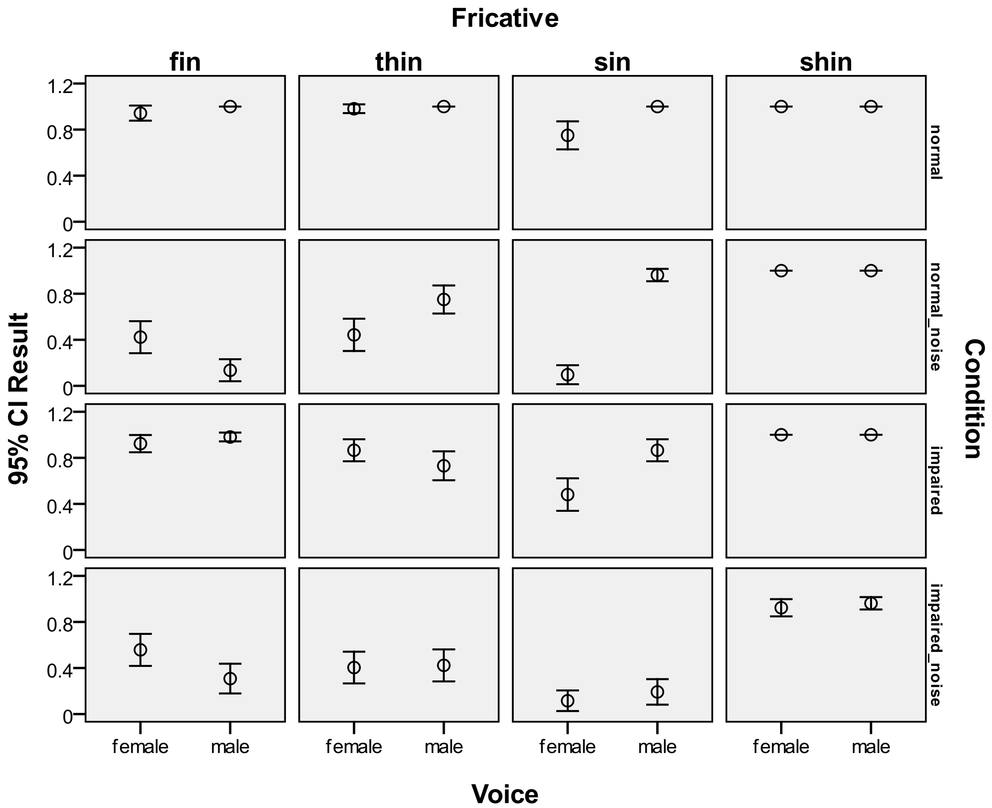


Figure 10 shows that in the no noise condition (normal, impaired), listeners made more mistakes in the female voice condition (13.2%) than in the male voice condition (5.3%). The same pattern can be observed in the noise condition (normal+noise and impaired+noise); more errors occurred in the female voice condition (50.5%) than in the male voice condition (40.9%).

A comparison of means revealed that the fricative /ʃ/ in *shin* (*M*=.99, SD=.119) was identified most accurately in both the male and female voice condition. The fricative /f/ (*M*=.66, SD=.475) and the fricative /θ/ (*M*=.70, SD=.459) received almost identical scores of correct recognition. The fricative /s/ (*M*=.56, SD=.497), on the other hand, received the lowest overall score. The Levene's test confirmed unequal variances [(F_{fin} = 8.587, p < .005); (F_{thin} = 2.212, p = .086); (F_{sin} = 9.161, p < .005); (F_{shin} = 2.691, p = .046)], indicating that not all listeners made the same errors. The listeners' answers differ most as regards the fricatives /s/ and /f/.

4.3.2 Analysis of reaction time (RT)

A Univariate Analysis of Variance was carried out to test the statistical significance of 'RT' on 'Results' (2). The results revealed a highly significant effect of RT on Results [F (1,1662)=75.976, p<0.005]. Figure 11 below shows how long it took a listener to recognise a fricative. On average, listeners were faster in opting for /s/ and /ʃ/ than for /f/. The slowest reaction time scores were recorded for the fricative /θ/ in both the experimental and the repetition session.

Figure 11: Impact of talker's voice on RT

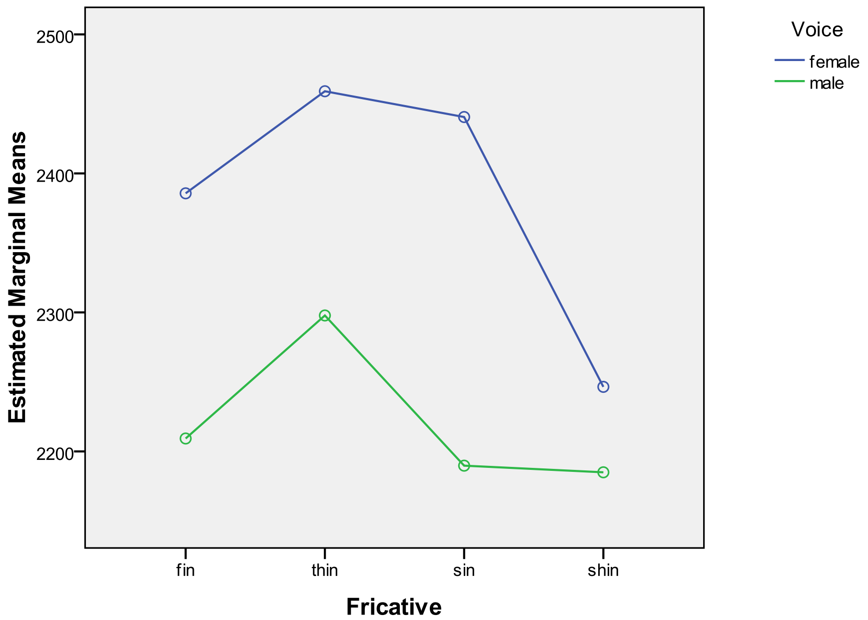


Figure 11 further illustrate that, on average, it took listeners more time to recognise fricatives in the female voice condition.

A One-sample *t*-test for 'Reaction Time (RT) and 'Fricative' also indicated that the fricatives influenced RT [($M_{/f/}$ = 2297.48 s, SD = 753.89 s), *t*(415) = 62.16, $p < 0.001$); ($M_{/θ/}$ = 2378.45 s, SD = 798.65 s, *t*(415) = 60.74, $p < 0.001$); $M_{/s/}$ = 2315.16 s, SD = 716.04 s, *t*(415) = 65.95, $p < 0.001$); ($M_{/ʃ/}$ = 2215.73 s, SD = 866.45 s, *t*(415) = 52.16, $p < 0.001$)]. The fastest mean RT was reported for the fricative /ʃ/. For the other fricatives, listeners needed more time to decide whether they heard /f/, /s/ or /θ/. Although /ʃ/ scored the fastest mean RT overall, variability within the group of listeners is greater than for the other fricatives. Furthermore, the results also showed that not only did /ʃ/ have the shortest RT, but also the most accurate results (99%). For /θ/, the mean recorded RT was highest and 70% of the answers were correct, compared to /f/ (66%) and /θ/ (70%) which had much faster RT.

Another aspect influencing RT was noise. A Univariate Analysis of Variance with 'Condition' (4), 'Fricative' (4) and 'Result' (2) revealed a highly significant effect on 'RT' [$F(6,1635) = 3.197$, $p > 0.005$).

Figure 12: Impact of noise on RT

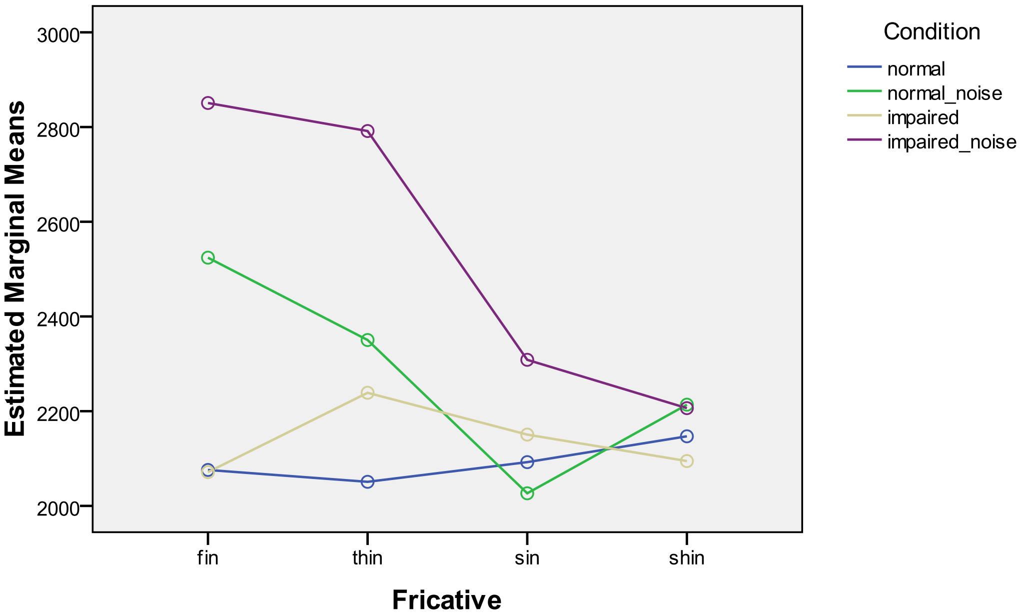


Figure 12 shows that most of the listeners had difficulties recognising fricatives in the impaired+noise condition and in the normal+noise condition. The noise conditions, however, did not considerably affect the recognition of the fricative /ʃ/.

4.3.3 Analysis of nationality

To test the statistical significance of Nationality on Results, a 3 x 2 One-Way ANOVA was conducted. The results reveal no significant effects of 'Nationality' on the recognition of fricatives [F (2,1663) =0.623, p=0.536].

6. Discussion

This study focused on the recognition of four English fricatives (/f/, /θ/, /s/, /ʃ/) and how our linguistic perception of fricatives can be affected by various auditory conditions. In the study, it was observed that the influence of noise interference had a greater effect on the recognition of fricatives than the impairment condition alone. The general results show that the non-sibilant fricatives /f/ and /θ/ were recognised better in the normal and impaired condition than the sibilant fricatives /s/ and /ʃ/. On the other hand, in the simulated hearing-impaired condition, sibilant fricatives are recognised better than non-sibilant fricatives. These results are congruent with Stevens et al. (1992) and Maniwa et al. (2008). Furthermore, the findings show that consonant confusion errors occurr most often in the noise condition, which confirms earlier findings by Miller and Nicely (1955), showing an almost identical pattern of consonant confusion for the four tested fricatives. However, it is important to note that there is considerable variability within the group of listeners regarding the correct recognition of fricatives, which affects in particular the results for /s/ and /ʃ/. This is to say that the palato-alveolar fricative /ʃ/ is recognised better than the alveolar fricative /s/ across all conditions.

An explanation for why the fricative /ʃ/ has been recognised better than all other fricatives across all conditions might due to its acoustic characteristics which help listeners distinguishing it form other fricatives, particularly from /s/. First, the frequency range for /ʃ/ is higher than for /s/. Second, /ʃ/ has a higher pitch than /s/ and is therefore much louder but has a shorter duration of frication noise than /s/. The longer duration of frication noise for /s/ conforms with Pirello et al. (1997). The fricative /ʃ/ also shows more energy in higher frequencies than other fricatives, even in the impaired condition, which makes it easier for listeners to recognise it. Another reason could be that masking is less effective for the high frequency energy of /ʃ/ so that most of the acoustic characteristics remain audible for listeners. In addition, the results for /s/ also confirm the findings by Skinner et al. (1997) who showed that /s/ is identified correctly in 60% of all cases even at the lowest sound level. In the present study, /s/ has also been recognised correctly (67.3%) in the hearing-impaired condition. Although this condition cannot be directly compared with a low sound level, it nevertheless illustrates that /s/ is recognised correctly even in distorted conditions.

Consonant confusion might have occurred due to the similar energy distribution of some of the fricatives. The acoustic spectrograms of the two fricatives /s/ and /f/ look very similar; that is to say the energy distribution only differs in the higher frequency region, which might have increased confusion in the listeners’ perception of the fricatives. Furthermore, duration measurements have shown that the average frication duration for /s/ and /ʃ/ tends to be longer than for /f/ and /θ/ (Behrens and Blumstein, 1988). In the present study, though, this is only true for the female voice recordings; a fact that might have led to greater consonant confusion. On the other hand, since /f/ and /θ/ are more frontal than /s/, they tend to be recognised better (Dubno and Levitt, 1981).

However, an explanation of why /f/, /θ/ and /s/ get confused so often in the noise condition might be that errors in noise conditions are assumed to be “a result of an increase in random guessing in the more difficult listening conditions”(Dubno and Levitt, 1981:254). The reason for this is that generally, English sentences are known to have rather low entropy, meaning that it is fairly predictable for a listener to guess what word follows another in a normal hearing condition. However, in the experiment, the listener could not depend on contextual information for sentence intelligibility, as the carrier sentence used in the study does not contain a lot of semantic information which could help a listener to recognise the fricative in the target word. Hence the listeners are bound to guess the fricative. Accordingly, Stilp and Kluender’s (2010) findings have shown that speech intelligibility is better predicted by non-linguistic sensory measures of uncertainty than by physical acoustic measures. This is to say that because our ears sometimes provide us with imperfect information, the brain needs to interpret that information through the process of phoneme restoration. Hence it is assumed that the strength of the illusion depends on the success and the speed of lexical access, suggesting that phonemic restoration heavily depends on the internal lexicon (Repp, 1991:148).

The recognition of fricatives in the impaired condition does not differ considerably from the normal condition across all fricatives. It might be that the impaired condition simulates telephone speech, meaning that merely frequency but not dB levels are affected. Hence the impaired recordings do not differ too much from normal speech. Furthermore, although the overall findings show better recognition of fricatives in the male voice condition than in the female voice condition, the difference might be explained by the fact that female voices are on average higher pitched than male voices and consequently more affected by the impairment condition with the band-pass filter than the male voice recordings.

Based on the perceptual results indicating that correct fricative recognition is greatly influenced by the noise condition, further analysis reveals that voice gender also played a role. There is a tendency for the fricatives to be identified better in the male voice condition than the female voice condition. These results, however, do not support the findings by Wang and Bilger (1973) showing that subjects made place of articulation confusions in the presence of noise, in particular with /s/-/ʃ/ when listening to a male voice. In contrast, the present study revealed that, in general, /f/-/θ/ gets confused more often than /s/-/ʃ/ in the male voice condition with noise. In the female voice condition, on the other hand, /s/ is recognised worst among all the fricatives and confused most often with other fricatives. Also, in the female voice condition, consonant confusion occurs on average more often with /s/-/ʃ/ than with /f/-

/θ/. One reason might be that in the present study, only two talkers were selected for the recordings of the carrier sentences; or the speech rhythm (Ferguson, 2004). Another factor influencing the results might have been the time of recording. Baker and Holding (1992) have investigated the impact of noise and speech on cognitive task performance and found that noise in the morning resulted in higher inaccuracies than noise in the afternoon. Furthermore, the sibilants (/ʃ/, /s/) in the male recording have considerably more acoustic energy than the non-sibilant fricatives (/f/, θ/), which is not so distinctly visible in the female voice recordings. The reason for it might be that the male and female recordings simply differ in too many respects from each other with regard to acoustic characteristics from such as speed of speech, speech quality etc.

The present findings also reveal that the slowest RT is recorded in the noise condition. This is in agreement with Baker and Holding (1992) who found that the slowest performance was recorded in the noise condition than in quiet and speech conditions. The results further show that noise has a far greater impact on RT than the impairment condition alone, thus supporting the results by Baker and Holding (1992). Moreover there is an observed tendency that the more time listeners needed to recognise fricatives, the more errors they made.

Nevertheless, the findings should be viewed in light of the limitations of the study. For example, the number of listeners and the number of talkers for the speech recordings might have influenced the results. Whereas a greater variety of speech samples might have influenced the recognition of fricatives, it would also have made the experiment last longer. Additionally, the listeners' auditory threshold was not tested nor was information about music consumption collected, which might have affected individual results. Furthermore, the effects of band-pass filtering on the recognition of fricatives in the present study are not clear; in particular as regards the difference in talkers' gender for the recognition of fricatives. In the present study, only pink noise was used for the hearing experiment. Hence generalisations about noise and performance cannot be made; in particular since the type of noise alone might have influenced the recognition of fricatives as have shown the findings by Parikh and Loizou (2005).

Hence, future studies might include technical modifications with regard to filter settings and noise adjustments for a more in-depth analysis of the acoustic data. To investigate hearing loss in the present target group, further data could be collected by testing older students as part of a longitudinal study. Consonant confusion could be studied further by including an analysis of vowel-consonant transitions as the intelligibility of fricatives also depends on how adjacent vowels retain information on preceding consonants. Furthermore, consonants in CV (consonant-vowel) clusters are more difficult to identify than VC (vowel-consonant) clusters. Studies have reported higher performances for consonants that are accompanied by the low back vowel /ɑ/ than for consonants that are paired with the high front vowel /i/ (Dubno and Levitt, 1981; Gordon-Salant, 1986). For this reason, a modified experimental design could include various CV and VC clusters including consonants accompanied by the vowels /ɑ/ or /u/ to compare the outcomes with the present findings. In addition, as an increasing number of young adults suffer from hearing loss; future studies might investigate how listeners use their own phonological resources in a hearing-impaired condition to process acoustic information in spoken utterances

7. Conclusion

The objective of the study was to investigate whether the same results as in previous studies (Dubno and Levitt, 1981; Maniwa et al. 2007; Miller and Nicely, 1955) can be replicated by testing students in a different language environment across several conditions. The findings showed that there are differences to be observed regarding the recognition of fricatives in the various conditions. Most of the errors in consonant confusion occurred in noise conditions. In general, the findings have shown that the fricative /ʃ/ is recognised best across all conditions, whereas /f/, /θ/ and /s/ had been confused fairly often. The present study also revealed that listeners confuse some of the fricatives even in the normal condition without noise interference or impaired signals. These findings might imply that students from the University of Edinburgh already encounter difficulties in distinguishing or hearing certain fricatives; whereas consonant confusion in the impaired condition might suggest that students also have difficulties recognising fricatives over the telephone. Young adults should therefore be made more aware of the detrimental and irreversible effects of noise-induced hearing loss. Fricatives in the English language constitute the largest set of consonants. Consequently, for a person suffering from high frequency hearing impairment it is more difficult to recognise certain consonants and also to fully capture the emotions that are carried through variation in pitch and amplitude, for instance. Even though some consonants tend to be louder than others and are therefore recognised better across various noise conditions, it does not change the fact that hearing loss is irreversible and only gets worse with age. The reason for it being that in the course of normal ageing, hair cells naturally break down and die. However, as the experimental setting was limited in time and size, the findings are approximations and only account for tendencies. Nevertheless, the results from the study do point to the necessity for a systematic study of noise-induced hearing loss in young adults.

References

Aithal, Sreedevi, Al Yonovitz and Venkatesh Aithal. (2008). Perceptual Consequences of Conductive Hearing Loss: Speech Perception in Indigenous Students Learning English as a ‘School’ Language. *The Australian and New Zealand Journal of Audiology,* 30:1, 1-18.

Badin, Pierre. (1989). Acoustics of voiceless fricatives, Production theory and data, *Speech Transmission Laboratory Quarterly Progress and Status Report, KTH,* 30:3, 33-55.

Baker, Mary Anne and Dennis H. Holding. (1992). The Effects of Noise and Speech on Cognitive Task Performance. *The Journal of General Psychology,* 120:3, 339-355.

Ben-David, Boaz M., Craig G. Chambers, Meredyth Daneman, M. Kathleen Pichora-Fuller, Eyal M. Reingold and Bruce A. Schneider. (2011). Effects of Aging and Noise on Real-Time Spoken Word Recognition: Evidence From Eye Movements. *Journal of Speech, Language, and Hearing Research.* 54, 243-262.

Behrens, Susan J. and Sheila E. Blumstein. (1988). Acoustic characteristics of English voiceless fricatives: a descriptive analysis. *Journal of Phoentics*, 16, 295-298.

Brinton, Laurel J. and Donna M. Brinton. (2010). *The Linguistic Structure of Modern English.* Amsterdam: John Benjamins B.V.

Cheung, Him, Hsuan-Chih Chen, Chun Yip Lai, On Chi Wong and Melanie Hills. (2001). The development of phonological awareness: effects of spoken language experience and orthography. *Cognition*. 81:3, 227-241.

Connine, Cynthia M., Dawn G. Blasko and Debra Titone. (1993). Do the Beginnings of Spoken Words Have a Special Status in Auditory Word Recognition? *Journal of Memory and Language*, 32, 193-210

Cutler, Anne, Andrea Weber, Roel Smits and Nicole Cooper. (2004). Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America,* 116:6, 3668–3678.

Darwin, Chris J. (1971). Ear differences in the recall of fricatives and vowels. *Quarterly Journal of Experimental Psychology*, 23, 46-62.

Dubno, Judy R. and Harry Levitt. (1981). Predicting consonant confusions from acoustic analysis. *The Journal of the Acoustical Society of America*, 69:1, 249-261.

Dubno, Judy R., Donald D. Dirks and Amy B. Schaefer. (1987). Effects of hearing loss on utilization of short-duration spectral cues in stop consonant recognition. *The Journal of the Acoustical Society of America*, 81:6, 1940-1947.

Ferguson, Sarah Hargus. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. *Journal of the Acoustical Society of America*, 116:4, 2365-2373.

Ferguson, Sarah Hargus and Diane Kewley-Port. (2010). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America,* 112:1, 259-271.

Freyman, Richard L., Uma Balakrishnan and Karen S. Helfer. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *Journal of the Acoustical Society of America,* 115:5, 2246–2256.

Friedrich, Gerhard, Wolfang Bigenzah and Patrick Zorowka (Eds). (2008). *Phoniatrie und Pädaudiologie, Einführung in die medizinischen, psychologischen und linguistischen Grundlagen von Stimme, Sprache und Gehör.* 4th Edition. Bern: Verlag Hans Huber

Gordon-Salant, Sandra. (1986). Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing. *The Journal of the Acoustical Society of America,* 80:6, 1599-1607.

Hayward, Katrina. (2000). *Experimental Phonetics.* Harlow: Pearson Education Limited.

Harley, A. Trevor. (2004). *The Psychology of Language. From Data to Theory.* 2nd Edition. East Sussex: Psychology Press.

Ingram, John C.L. (2007). *Neurolinguistics, An Introduction to Spoken Language Processing and Its Disorders.* Cambridge: Cambridge University Press.

Johnson, Keith. (2004). *Acoustics and Auditory Phonetics*. 2nd Edition. Blackwell Publishing: Malden.

Kennedy, Elizabeth, Harry Levitt, Arlene C. Neuman and Mark Weiss. (1998). *The Journal of the Acoustical Society of America*, 103:2, 1098-1114.

Kompis, Martin. (2009). *Audiologie.* 2nd Edition. Bern: Verlag Hans Huber

Ladefoged, Peter. (2003). *Phonetic Data Analysis. An Introduction to Fieldwork and Instrumental Techniques.* Malden: Blackwell Publishing.

Ladefoged, Peter. (2005). *Vowels and Consonants.* 2nd edition. Malden: Blackwell Publishing.

Lipscomb, D. M. (1972). The Increase in Prevalence of High Frequency Hearing Impairment among College Students. *Audiology*, 11, 231-237.

Le Prell, Colleen, G., Brittany N. Hensley, Kathleen C. Campbell, James W. Hall III and Kenneth Guire K. (2011). Evidence of hearing loss in a “normally-hearing” college student population. *International Journal of Audiology,* 50: Supplement 1, 21-31.

Magrab, Edward B. (1975). *Environmental Noise Control.* New York: John Wiley & Sons.

Mackersie, Carol L., James Dewey and Lesli A. Guthrie. (2011). Effects of fundamental frequency and vocal-tract length cues on sentence segregation by listeners with hearing loss. *Journal of the Acoustical Society of America,* 130:3, 1006-1019.

Maniwa, Kazumi, Allard Jongman and Travis Wade. (2008). Perception of clear fricatives by normal-hearing and simulated hearing-impaired listeners. *The Journal of the Acoustical Society of America,* 123:2, 1114-1125.

Maniwa, Kazumi, Allard Jongman and Travis Wade. (2009). Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America,* 125:6, 3962–3973.

Miller, George A. and Patricia E. Nicely. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338–352.

Nishi, Kanae, Dawna E. Lewis, Brenda M. Hoover, Sandgsook Choi and Patricia G. Stelmachowicz. (2010). Children's recognition of American English consonants in noise. *The Journal of the Acoustical Society of America,* 127:5, 3177-3188.

Panda, Naresh K., Sanjay Munjal and Jaimanti Kakshi. (2010). Audiological Disturbances in Long-Term Mobile Phone Users. *Journal of Otolaryngol Head Neck Surgery*, 39:1, 5 – 11.

Parikh, Gaurang and Philipos C. Loizou (2005). The influence of noise on vowel and consonant cues. *Journal of the Acoustical Society of America,* 118:6, 3874–3888

Pearson, Jay D., Christopher H. Morrell, Sandra Gordon-Salant, Larry J. Brant, E. Jeffrey Metter, Lisa L. Klein and James L. Fozard. (1995). Gender differences in a longitudinal study of age-related hearing loss. *The Journal of the Acoustical Society of America,* 197:2, 1196-1205..

Pirello, Karen, Sheila E. Blumstein and Kathleen Kurowski. (1997). The characteristics of voicing in syllable-initial fricatives in American English. *The Journal of the Acoustical Society of America,* 101:6, 3754-3765.

Rabinowitz, Peter M. (2000). Noise-Induced Hearing Loss. *AAFP*. Retrieved on 25 July 2011 from *http://www.aafp.org/afp/20000501/2749.html*.

Repp, Bruno H. (1991). Perceptual Restoration of a "Missing" Speech Sound: Auditory Induction or Illusion? *Haskins Laboratories. Status Report on Speech Research,* SR 107/108, 147-170.

Rosen, Stuart. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society: Biological Sciences*, 336:1278, 367-373.

Ross, Mark. (2005). Frequency-Lowering Hearing Aids: Increasing the Audibility of High-Frequency Speech Sounds. *Rehabilitation Engineering Research Center on Hearing Enhancement (RERC-HE).* Retrieved on 25 July 2011 from http://www.hearingresearch.org/ross/hearing_aid_use/frequency-lowering_hearing_aids.php.

Rubin, Philip, M. T. Turvey and Peter van Gelder. (1976). Initial phonemes are detected faster in spoken words than in spoken nonwords. *Perception and Psychophysics,* 19:5, 394-398.

Schneider, Bruce A., Meredyth Daneman and Dana R. Murphy. (2005). Speech comprehension difficulties in older adults: Cognitive slowing or age-related changes in hearing? *Psychology and Aging,* 20:2, 261-271.

Shargorodsky, Josef, Sharon G. Curhan, Gary C. Curhan and Roland Eavey (2010). Change in Prevalence of Hearing Loss in US Adolescents. *JAMA*, 304:7, 772-778.

Simpson, Adrian P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass,* 3:2, 621-640.

Skinner, Margaret W., Laura K. Holden, Timothy A. Holden, Marilyn E. Deorest and Marios S. Fourakis. (1997). Speech recognition at simulated soft, conversational, and raised-to-loud vocal efforts by adults with cochlear implants. *Journal of the Acoustical Society of America,* 101:6, 3766-3782.

Stevens, Kenneth N., Sheila E. Blumstein, Laura Glicksman, Martha Burton and Kathleen Kurowski. (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *The Journal of the Acoustical Society of America*, 91:5, 2979-3000.

Stevens, Kenneth N. (1997). Articulatory-Acoustic-Auditory Relationships. In: William J. Hardcastle and John Laver, eds., *The Handbook of Phonetics Sciences,* 462-506. Oxford: Blackwell.

Stilp, Christian E., Michael Kiefte, Joshua M. Alexander and Keith R. Kluender. (2010). Cochlea-scaled spectral entropy predicts rate-invariant intelligibility of temporally distorted sentences. *The Journal of the Acoustical Society of America*, 128:4, 2112–2126.

Stilp, Christian E. and Keith R. Kluender. (2010). Cochlea-scaled entropy, not consonants, vowels, or time, best predicts speech intelligibility. *PNAS*, 107:27, 12387-12392.

Strevens, Peter. (1960). Spectra of fricative noise in human speech. *Language and Speech,* 3, 32-49.

Vogel, Ineke, Hans Verschuure, Catharina P. B. van der Ploeg, Johannes Brug and Hein Raat. (2010). Estimating Adolescent Risk for Hearing Loss Based on Data From a Large School-Based Survey**.** *American Journal of Public Health*, 100:6, 1095-1100.

Vogel, Ineke, Johannes Brug, Catharina P.B. van der Ploeg and Hein Raat (2009). Strategies for the Prevention of MP3-Induced Hearing Loss Among Adolescents: Expert Opinions From a Delphi Study. *Petriatrics*, 123, 1257-1262.

Wang, Marilyn D. and Robert C. Bilger. (1973). Consonant confusion in noise: A study of perceptual features. *The Journal of the Acoustical Society of America,* 54:5, 1248-1266.

Zeng, Fan-Gen and Christopher W. Turner. (1990). Recognition of Voiceless Fricatives by Normal and Hearing-Impaired Subjects. *Journal of Speech and Hearing Research,* 33, 440-449.